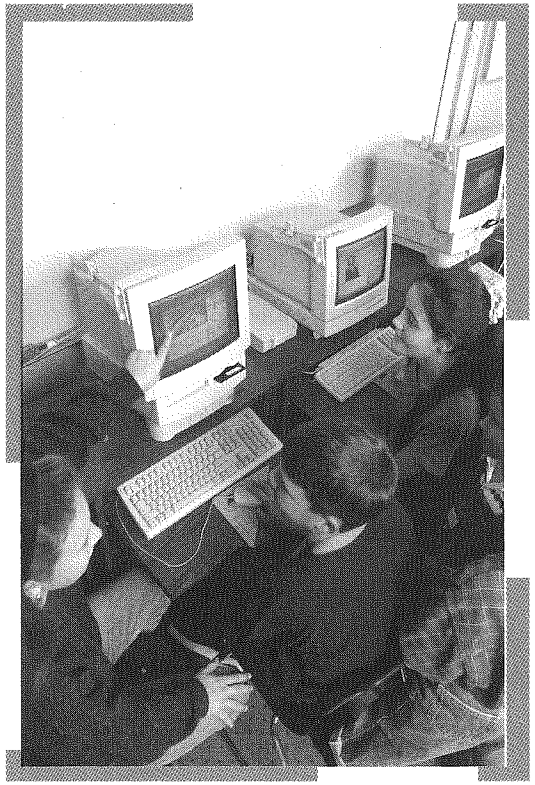
Ο Δημήτρης Καραγεώργος είναι μαθηματικός με διδακτορικές σπουδές στην Αγγλία. Υπηρέτησε τριάντα πέντε χρόνια στη Δευτεροβάθμια Εκπαίδευση ως καθηγητής, γυμνασιάρχης, λυκειάρχης, σχολικός σύμβουλος, μόνιμος πάρεδρος και σύμβουλος του Παιδαγωγικού Ινστιτούτου.
Σήμερα διδάσκει ως επίκουρος καθηγητής στο Πανεπιστήμιο Αθηνών (τομέας Παιδαγωγικής του τμήματος Φ.Π.Ψ.) τα μαθήματα «Διδακτική των θετικών επιστημών» και «Μεθοδολογία έρευνας στις επιστήμες της αγωγής».
Στη μακρόχρονη εκπαιδευτική του θητεία δίδαξε το μάθημα των μαθηματικών σε όλες τις τάξεις του Γυμνασίου και Λυκείου, ασχολήθηκε με τη σύνταξη αναλυτικών προγραμμάτων, τη συγγραφή διδακτικών εγχειριδίων, που διδάσκονται στη Δευτεροβάθμια Εκπαίδευση, την παραγωγή υποστηρικτικού διδακτικού υλικού για τα μαθηματικά και την επιμόρφωση των εκπαιδευτικών.
Έχει γράψει εννιά βιβλία μαθηματικών και έχει δημοσιεύσει σε ελληνικά και ξένα περιοδικά περισσότερες από πενήντα εργασίες που αναφέρονται στα προηγούμενα αντικείμενα και στον εκπαιδευτικό σχεδιασμό.
Ένα από τα θέματα που τον απασχόλησαν την τελευταία εικοσαετία ήταν και η εκπαιδευτική έρευνα.
Ένα μέρος της προσπάθειάς του αυτής καταγράφεται σε αυτό το βιβλίο.

Δημήτρης Λ. Καραγεώργος

ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΑΘΗΝΩΝ
Επιτ. ΣΥΜΒΟΥΛΟΣ ΤΟΥ ΠΑΙΔΑΓΩΓΙΚΟΥ ΙΝΣΤΙΤΟΥΤΟΥ

ΜΕΘΟΔΟΛΟΓΙΑ ΕΡΕΥΝΑΣ
στις επιστήμες της αγωγής

ΜΙΑ ΔΙΔΑΚΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ



Σαββάλας
ΕΚΔΟΣΕΙΣ

πει να δούμε ότι η αποτελεσματικότητα της μεθόδου, σε αυτή την περίπτωση, εξαρτάται από το μέγεθος της τάξης στην οποία χρησιμοποιείται. Αυτό φανερώνει ότι υπάρχει αλληλεπίδραση μεταξύ μεθόδων και μεγέθους της τάξης.

Αναφέραμε στα προηγούμενα ότι, προκειμένου να διαπιστώσουμε τη σημαντικότητα των διαφορών που παρατηρούνται στην εφαρμογή κάποιας θεραπείας (διδακτικής μεθόδου) στην πειραματική ομάδα και στη μη εφαρμογή στην ομάδα ελέγχου, καταφεύγουμε σε ορισμένα στατιστικά κριτήρια. Θα αναφέρουμε εδώ δύο από αυτά. Το x^2 - κριτήριο και την ανάλυση της διασποράς (ANOVA).

6.3 Έλεγχος υποθέσεων με τη X^2 - κατανομή Μη παραμετρικά στατιστικά

Τα περισσότερα από τα στατιστικά τεστ αναφέρονται στις παραμέτρους πληθυσμών (μέσους, διακυμάνσεις, ποσοστά κ.λπ.) και γι' αυτό τα ονομάζουμε **παραμετρικά τεστ**.

Πολλές φορές όμως οι ερευνητές αντιμετωπίζουν πειραματικές καταστάσεις οι οποίες δεν υπακούουν (δεν μπορούν να ενταχθούν) στις απαιτήσεις των παραμετρικών τεστ. Γι' αυτές τις περιπτώσεις έχουν αναπτυχθεί τεχνικές ελέγχου υποθέσεων που δεν ακολουθούν τους κανόνες των παραμετρικών ελέγχων. Αυτές οι εναλλακτικές λύσεις ονομάζονται **μη παραμετρικά τεστ**.

Θα διαπιστώσουμε ότι αυτά τα μη παραμετρικά τεστ συνήθως δεν δηλώνουν υποθέσεις με όρους παραμέτρων του πληθυσμού, αλλά κάνουν μερικές υποθέσεις για την κατανομή του πληθυσμού. Άλλη μια διάκριση των μη παραμετρικών τεστ είναι ότι χρησιμοποιούνται σε δεδομένα που μετρούνται σε ονομαστικές ή τακτικές κλίμακες. Σημειώνουμε ότι τα μη παραμετρικά τεστ δεν είναι τόσο ευαίσθητα όσο τα παραμετρικά τεστ. Μη παραμετρικά τεστ είναι πιο πιθανό να αποτύχουν να ανακαλύψουν διαφορές μεταξύ δύο διδακτικών μεθόδων, θεραπειών κ.λπ. Γι' αυτό, **όταν έχουμε επιλογή μεταξύ παραμετρικών και μη παραμετρικών τεστ, σαφώς θα προτιμούμε τα παραμετρικά**.

Γενικά, η X^2 - κατανομή χρησιμοποιείται κυρίως για τον έλεγχο στατιστικών υποθέσεων οι οποίες στηρίζονται στις διαφορές μεταξύ των **παρατηρούμενων συχνοτήτων** (observed frequencies - f_o) και των αντίστοιχων **αναμενόμενων συχνοτήτων** (expected frequencies - f_e).

Για να εξάγουμε συμπεράσματα με το κριτήριο x^2, συγκρίνουμε τις πραγματικές (παρατηρούμενες) (f_o) συχνότητες με τις αντίστοιχες αναμενόμενες (f_e) ή θεωρητικές συχνότητες. Δηλαδή δεν προσδιορίζουμε καμία παράμετρο του πληθυσμού και γι' αυτό η X^2 - κατανομή ονομάζεται και **μη παραμετρικό κριτήριο ελέγχου** (non - parametric test). Αυτό απλά σημαίνει ότι δεν μας ενδιαφέρει η κανονικότητα ή μη

του πληθυσμού από τον οποίο παίρνουμε τα δείγματα, σε αντίθεση με τα κριτήρια z και t(v) που λέγονται παραμετρικά, γιατί ακριβώς προσδιορίζουμε παραμέτρους των πληθυσμών και προϋποθέτουμε κανονικότητα πληθυσμών από τους οποίους προέρχονται τα δείγματα.

Σημειώνουμε ότι για να είναι έγκυρο το κριτήριο x^2, πρέπει το μέγεθος του δείγματος να είναι τουλάχιστον 50 και οι επιμέρους συχνότητες τουλάχιστον 5. Ας δούμε μ' ένα παράδειγμα πώς κατασκευάζουμε τη X^2 - κατανομή και τη χρήση της σε ορισμένες περιπτώσεις.

Α. Η X^2 - κατανομή.

Έστω ότι ρίχνουμε ένα ζάρι 120 φορές και καταγράφουμε τις εμφανίσεις των σημείων 1, 2, 3, 4, 5 και 6. Η κατανομή φαίνεται στον ακόλουθο πίνακα.

Αριθμός σημείων	1	2	3	4	5	6
Πραγματικές συχνότητες	16	19	27	17	23	18

Αν το ζάρι ήταν «καθαρό» (τίμιο), τότε τα αναμενόμενα (θεωρητικά) ποσοστά σε καθεμία περίπτωση θα ήταν $\frac{1}{6}$, που είναι ισοδύναμο με το να πούμε ότι η πιθανότητα να εμφανιστεί μια περίπτωση, π.χ. το 2, είναι $\frac{1}{6}$. Έτσι η αναμενόμενη συχνότητα σε κάθε περίπτωση θα ήταν 20, οπότε μπορούμε να έχουμε τον ακόλουθο πίνακα με τις πραγματικές και τις θεωρητικές συχνότητες.

Αριθμός σημείων	1	2	3	4	5	6
Πραγματικές συχνότητες (f_o)	16	19	27	17	23	18
Αναμενόμενες συχνότητες (f_e)	20	20	20	20	20	20

f_o = observed frequencies

f_e = expected frequencies

Ένας τρόπος να μετρήσουμε τις διαφορές μεταξύ πραγματικών και αναμενόμενων συχνοτήτων είναι απλά να αφαιρέσουμε τις αναμενόμενες από τις πραγματικές συχνότητες. Έτσι έχουμε:

Αριθμός σημείων	1	2	3	4	5	6
$f_o - f_e$	-4	-1	7	-3	3	-2

Παρατηρούμε ότι $\sum(f_o - f_e) = 0$ και αυτό συμβαίνει πάντα. Επομένως το άθροισμα των διαφορών δεν είναι ικανοποιητικό μέτρο. Έτσι μπορούμε να πάρουμε το $\sum(f_o - f_e)^2 = (-4)^2 + (-1)^2 + 7^2 + (-3)^2 + 3^2 + (-2)^2 = 88$ και αυτό είναι ένα ικανοποιητικό μέτρο των διαφορών $f_o - f_e$.

Αν όμως πάρουμε τους επόμενους πίνακες δεδομένων:

f_o	16	12	7
f_e	9	15	11

f_o	457	297	496
f_e	450	300	500

θα δούμε ότι και στις δύο περιπτώσεις έχουμε $\sum(f_o - f_e)^2 = 74$, αλλά οι διαφορές $f_o - f_e$ στον ένα πίνακα έχουν τελείως διαφορετική σημαντικότητα απ' ό,τι στον άλλο. Για παράδειγμα, η $f_o = 16$ είναι μεγαλύτερη σχεδόν κατά 78% της αντίστοιχης $f_e = 9$, ενώ η διαφορά των 7 μονάδων (457 − 450) είναι μόλις το 2% του 450. Γι' αυτό παίρνουμε ένα καλύτερο μέτρο, το οποίο είναι

$$x^2 = \sum \frac{(f_o - f_e)^2}{f_e} \qquad (1)$$

Για το παράδειγμα του ζαριού θα έχουμε:

$$x^2 = \sum \frac{(f_o - f_e)^2}{f_e} =$$

$$= \frac{(16-20)^2}{20} + \frac{(19-20)^2}{20} + \frac{(27-20)^2}{20} + \frac{(17-20)^2}{20} + \frac{(23-20)^2}{20} + \frac{(18-20)^2}{20} = 4{,}4$$

Ο τύπος (1), κάνοντας πράξεις, γράφεται:

$$x^2 = \sum \frac{f_o^2 - 2f_of_e + f_e^2}{f_e} = \sum \frac{f_o^2}{f_e} - \sum 2f_o + \sum f_e =$$

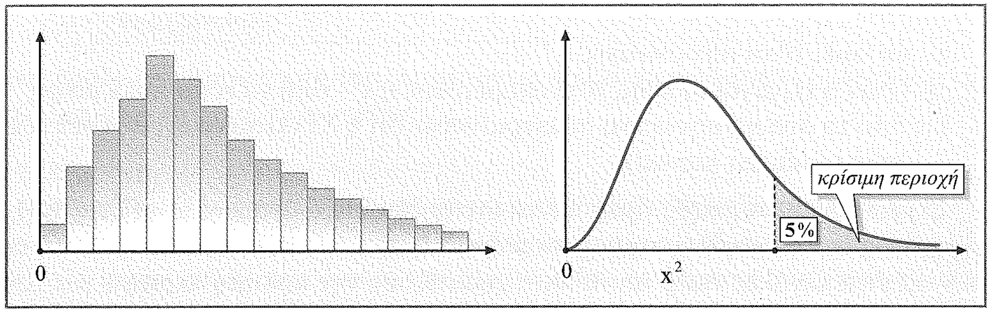
$$= \sum \frac{f_o^2}{f_e} - 2\sum f_o + \sum f_e = \sum \frac{f_o^2}{f_e} - 2\nu + \nu = \sum \frac{f_o^2}{f_e} - \nu$$

αφού $\sum f_o = \sum f_e = 120 = \nu$.

Είναι φανερό ότι αν επαναλάβουμε το πείραμα πολλές φορές (ρίχνουμε το ζάρι 120 φορές και καταγράφουμε τις εμφανίσεις των σημείων - αριθμών 1, 2, 3, 4, 5 και 6), μπορούμε να έχουμε διάφορα αποτελέσματα κάθε φορά. Για παράδειγμα, να έρθει 120 φορές 1, 0 φορές 1 ή 120 φορές 6 κ.λπ. ή σ' ένα πείραμα να έχουμε 20 φορές 1, 20 φορές 2, ..., 20 φορές 6, δηλαδή $f_o - f_e = 0$ για όλες τις εμφανίσεις. Αν υπολογίσουμε το $x^2 = \sum \frac{(f_o - f_e)^2}{f_e}$ για όλους τους δυνατούς συνδυασμούς των πραγματικών συχνοτήτων (γιατί οι αναμενόμενες είναι πάντα 20 για κάθε εμφάνιση), θα έχουμε έναν μεγάλο αριθμό τιμών για τη x^2.

Αν οι x^2 - τιμές ταξινομηθούν και κατασκευαστεί το ιστόγραμμα της κατανομής συχνοτήτων, τότε θα δούμε ότι θα μοιάζει με το ακόλουθο ιστόγραμμα ή με την

αντίστοιχη γραφική παράσταση, όταν οι δοκιμές (το πείραμα) αυξηθούν υπερβολικά.



Ας υποθέσουμε ότι το 5% του εμβαδού που περικλείεται από τη X^2 - καμπύλη είναι δεξιά της τιμής 11,07. Τότε η υπόθεση ότι το ζάρι είναι «καθαρό» θα απορριφθεί όταν η τιμή του x^2 που θα υπολογιστεί από ένα δείγμα (από ένα πείραμα) είναι μεγαλύτερη του 11,07.

Η X^2 - κατανομή, της οποίας η γραφική παράσταση εξαρτάται από τους βαθμούς ελευθερίας, χρησιμοποιείται ευρύτατα όπως είδαμε στα προηγούμενα και όπως θα δούμε στη συνέχεια.

Ο γενικός σκοπός σε κάθε έλεγχο υποθέσεων είναι να προσδιορίσουμε αν τα δεδομένα ενός δείγματος υποστηρίζουν ή διαψεύδουν μια υπόθεση για τον πληθυσμό. Σ' ένα x^2 - τεστ, για καλή προσαρμογή, το δείγμα εκφράζεται από το σύνολο των πραγματικών συχνοτήτων (f_o τιμές) και η μηδενική υπόθεση χρησιμοποιείται για να δημιουργηθεί ένα σύνολο αναμενόμενων συχνοτήτων (f_e τιμές).

Το x^2 - στατιστικό απλά μετρά πόσο καλά τα δεδομένα (f_o) προσεγγίζουν τα αναμενόμενα (f_e).

Ο τύπος για το x^2 (chi - square) - στατιστικό είναι $x^2 = \sum \frac{(f_o - f_e)^2}{f_e}$.

Β. Η X^2 - κατανομή και οι βαθμοί ελευθερίας.

Πρέπει να γίνει σαφές από τον τύπο του στατιστικού x^2 ότι η τιμή του μετρά τη διαφορά μεταξύ των πραγματικών συχνοτήτων (δεδομένα) και των αναμενόμενων συχνοτήτων (f_e). Όταν υπάρχουν μεγάλες διαφορές μεταξύ f_o και f_e, η τιμή του x^2 είναι μεγάλη και συμπεραίνουμε ότι τα δεδομένα δεν προσεγγίζουν την υπόθεση. Δηλαδή μια μεγάλη τιμή του x^2 οδηγεί στην απόρριψη της H_0. Τα αντίθετα ακριβώς συμβαίνουν όταν οι διαφορές $f_o - f_e$ είναι μικρές. Δηλαδή μια μικρή τιμή του x^2 δηλώνει ότι αποτυγχάνουμε να απορρίψουμε την H_0 (δεχόμαστε την H_0).

Για να αποφασίσουμε αν μια συγκεκριμένη τιμή του x^2 είναι «μεγάλη» ή «μικρή», πρέπει να αναφερθούμε σε μια X^2 - κατανομή. Αυτή η κατανομή είναι το σύνολο των τιμών του x^2 για όλα τα πιθανά τυχαία δείγματα όταν η H_0 είναι αληθής. Όπως και οι άλλες κατανομές (t, F κ.λπ.), η X^2 - κατανομή είναι μια θεωρητική κατανομή με καλά ορισμένα χαρακτηριστικά. Μερικά από αυτά είναι τα ακόλουθα:

i) Οι τιμές είναι μη αρνητικές.
ii) Όταν η H_0 είναι αληθής, τότε αναμένουμε η x^2 - τιμή να είναι πολύ μικρή.
iii) Από τα δύο προηγούμενα συμπεραίνουμε ότι η X^2 - κατανομή είναι θετικώς λοξή (positively skewed).

Σημειώνουμε ότι μικρές τιμές, κοντά στο μηδέν, είναι αναμενόμενες όταν η H_0 είναι αληθής και μεγάλες τιμές (στο άκρο δεξιά) είναι πολύ απίθανες. Έτσι, οι ασυνήθιστα μεγάλες τιμές του x^2 θα σχηματίζουν την κρίσιμη περιοχή για τον έλεγχο υποθέσεων.

Είπαμε ότι η καμπύλη της X^2-κατανομής είναι θετικώς λοξή, αλλά για να μιλήσουμε για το ακριβές σχήμα της, πρέπει να λάβουμε υπόψη μας τον αριθμό των κατηγοριών. Είναι φανερό ότι όσο περισσότερες κατηγορίες έχουμε, τόσο το άθροισμα $\sum \frac{(f_o - f_e)^2}{f_e}$ θα είναι μεγαλύτερο. Το αποτέλεσμα είναι ότι έχουμε μια οικογένεια X^2 - κατανομών, που καθεμία έχει το σχήμα της, το οποίο προσδιορίζεται από τον αριθμό των κατηγοριών που μελετάμε. Οι βαθμοί ελευθερίας (degrees of freedom) είναι $c - 1$, όπου c είναι ο αριθμός των κατηγοριών.

Σημειώνουμε με έμφαση ότι για ένα x^2 - τεστ οι βαθμοί ελευθερίας δεν έχουν καμία σχέση με το μέγεθος του δείγματος, όπως είναι σε όλα τα άλλα τεστ, αλλά εξαρτώνται από τον αριθμό των κατηγοριών, γιατί μπορεί μερικές συχνότητες να είναι μικρότερες του 5 και να τις αθροίσουμε (συνδυάσουμε) με άλλες.

Για να προσδιορίσουμε αν μια συγκεκριμένη x^2 - τιμή είναι εξαιρετικά μεγάλη, πρέπει πρώτα να καθορίσουμε το α - επίπεδο (0,05 ή 0,01) και έπειτα να συμβουλευτούμε τον πίνακα της X^2 - κατανομής, που βρίσκεται στους πίνακες του παραρτήματος.

Παράδειγμα

Ένας ερευνητής ενδιαφέρεται για τους παράγοντες που εμπλέκονται στην επιλογή των μαθημάτων από τους φοιτητές. Γι' αυτό επιλέγει τυχαία 50 φοιτητές και τους ζητά να απαντήσουν στην ερώτηση: «Ποιον από τους ακόλουθους παράγοντες θεωρείτε τον σπουδαιότερο όταν επιλέγετε ένα μάθημα;».

Α. Το ενδιαφέρον σας για το μάθημα.
Β. Την ευκολία του μαθήματος για να το περάσετε.
Γ. Τον καθηγητή που διδάσκει το μάθημα.
Δ. Το ωράριο που διδάσκεται το μάθημα κάθε εβδομάδα.

Ο κάθε ερωτώμενος οφείλει να επιλέξει έναν μόνο παράγοντα. Οι απαντήσεις των φοιτητών φαίνονται στον ακόλουθο πίνακα.

Παράγοντες	Α	Β	Γ	Δ
Συχνότητα προτίμησης (πραγματικές)	18	17	7	8

Να βρεθεί αν κάποιος από τους παράγοντες παίζει σπουδαιότερο ρόλο από τους άλλους στην επιλογή ενός μαθήματος.

Απάντηση

Βήμα 1ο: Δηλώνουμε τις υποθέσεις και το επίπεδο σημαντικότητας α.

H_0: Ο πληθυσμός των φοιτητών δεν δείχνει καμία προτίμηση επιλέγοντας κάποιον παράγοντα έναντι των άλλων. Έτσι, οι τέσσερις παράγοντες έχουν την ίδια πιθανότητα να επιλεγούν και η κατανομή του πληθυσμού έχει τις ακόλουθες πιθανότητες (ή ποσοστά):

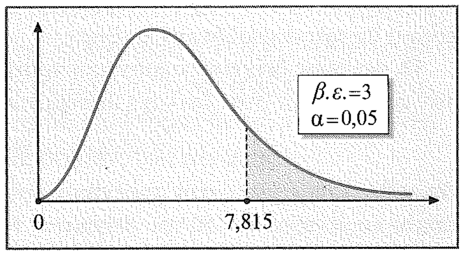
Παράγοντες	Α	Β	Γ	Δ
H_0: Ποσοστά (πιθανότητες)	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

H_1: Στον πληθυσμό των φοιτητών ένας ή περισσότεροι από αυτούς τους παράγοντες παίζουν μεγαλύτερο ρόλο στην επιλογή μαθήματος (δηλαδή κάποιος παράγοντας σημειώνεται περισσότερες φορές από τους φοιτητές).

Το επίπεδο σημαντικότητας θέλουμε να είναι (σταθερό συνήθως) $\alpha = 0{,}05$.

Βήμα 2ο: Προσδιορίζουμε τους βαθμούς ελευθερίας και έπειτα την κρίσιμη περιοχή. Εδώ οι βαθμοί ελευθερίας (β.ε.) είναι $c - 1 = 4 - 1 = 3$.

Για 3 βαθμούς ελευθερίας και $\alpha = 0{,}05$ μας δίνεται $x^2_{0,05}(3) = 7{,}815$ από τον πίνακα της X^2 - κατανομής.



Βήμα 3ο: Πρέπει να προσδιοριστούν οι αναμενόμενες συχνότητες για όλες τις κατηγορίες και έπειτα να υπολογιστεί το x^2 - στατιστικό.

Αν η H_0 ήταν αληθής και οι φοιτητές δήλωναν χωρίς προτίμηση σε κάποιον παράγοντα, τότε το ποσοστό των απαντήσεων σε κάθε παράγοντα θα ήταν 25%, δηλαδή $p = \frac{1}{4}$. Κι επειδή το μέγεθος του δείγματος είναι 50, η μηδενική υπόθεση προβλέπει αναμενόμενες συχνότητες $f_e = p \cdot v = \frac{1}{4} \cdot 50 = 12{,}5$ για όλες τις κατηγορίες. Έτσι μπορούμε να έχουμε τον ακόλουθο πίνακα πραγματικών και αναμενόμενων συχνοτήτων.

Παράγοντες	Α	Β	Γ	Δ
Πραγματικές συχνότητες (f_o)	18	17	7	8
Αναμενόμενες συχνότητες (f_e)	12,5	12,5	12,5	12,5

Χρησιμοποιώντας τις τιμές αυτού του πίνακα μπορούμε να υπολογίσουμε το x^2 - στατιστικό από τον τύπο $x^2 = \sum \frac{(f_o - f_e)^2}{f_e}$, δηλαδή

$$x^2 = \frac{(18 - 12{,}5)^2}{12{,}5} + \frac{(17 - 12{,}5)^2}{12{,}5} + \frac{(7 - 12{,}5)^2}{12{,}5} + \frac{(8 - 12{,}5)^2}{12{,}5} \Leftrightarrow$$

$$\Leftrightarrow x^2 = \frac{30{,}25}{12{,}5} + \frac{20{,}25}{12{,}5} + \frac{30{,}25}{12{,}5} + \frac{20{,}25}{12{,}5} =$$

$$= 2{,}42 + 1{,}62 + 2{,}42 + 1{,}62 = 8{,}08$$

Βήμα 4ο: Η τιμή 8,08 του στατιστικού πέφτει μέσα στην κρίσιμη περιοχή, αφού $8{,}08 > 7{,}815$. Έτσι απορρίπτουμε την H_0 και ο ερευνητής μπορεί να συμπεράνει πως οι φοιτητές δηλώνουν ότι μερικοί παράγοντες είναι περισσότερο σημαντικοί από τους άλλους για την επιλογή των μαθημάτων που θα παρακολουθήσουν.

Σημείωση: Αν αυτή η έρευνα δημοσιευόταν σε κάποιο επιστημονικό περιοδικό, το συμπέρασμα θα παρουσιαζόταν ως εξής:

Οι φοιτητές δείχνουν σημαντική προτίμηση στην ερώτηση που αφορά τους παράγοντες που εμπλέκονται στην επιλογή μαθημάτων

$$x^2(3, v = 50) = 8{,}08, \quad p < 0{,}05$$

Όπως βλέπουμε, μέσα στην παρένθεση πρώτα γράφονται οι βαθμοί ελευθερίας (β.ε. = 3) και έπειτα το μέγεθος $v = 50$ του δείγματος των φοιτητών. Υπενθυμίζουμε ότι το $p < 0{,}05$ δηλώνει ότι η πιθανότητα να κάνουμε σφάλμα τύπου Ι (δηλαδή να απορρίψουμε την H_0, ενώ είναι αληθής) είναι μικρότερη του $\alpha = 0{,}05$.

Γ. Έλεγχος ανεξαρτησίας ιδιοτήτων - Πίνακες συσχέτισης (συνάφειας).

Το κριτήριο x^2 μπορεί να χρησιμοποιηθεί για τον έλεγχο της ανεξαρτησίας ή όχι μεταξύ δύο ιδιοτήτων. Υπάρχουν δηλαδή περιπτώσεις κατά τις οποίες ένα δείγμα ταξινομείται ανάλογα με δύο **ποιοτικές** ιδιότητες ή με **ποσοτικά** χαρακτηριστικά και θέλουμε να ελέγξουμε αν η μια ιδιότητα έχει κάποια επίδραση στην άλλη. Για να κάνουμε αυτόν τον έλεγχο, κατατάσσουμε τα δεδομένα σε πίνακα διπλής εισόδου με **r γραμμές** και **c στήλες**. Ο πίνακας αυτός λέγεται **πίνακας συσχέτισης** ή **συνάφειας**.

Εδώ η H_0 θα εκφράζει την **ανεξαρτησία** των ιδιοτήτων, δηλαδή ότι η μια ιδιότητα δεν έχει καμία επίδραση στην άλλη, και η H_1 ακριβώς το αντίθετο. Με βάση την υπόθεση της ανεξαρτησίας, υπολογίζουμε τις αναμενόμενες (θεωρητικές) συχνότητες f_e, οπότε παίρνουμε το στατιστικό $x^2 = \sum \frac{(f_o - f_e)^2}{f_e}$ με βαθμούς ελευθερίας β.ε. $= (r - 1)(c - 1)$, για να κάνουμε τον έλεγχο της υπόθεσης.

Αν $x^2 < x^2_\alpha$ (β.ε.), όπου α το επίπεδο σημαντικότητας (συνήθως $\alpha = 0{,}05$), τότε αποδεχόμαστε την H_0, δηλαδή την ανεξαρτησία των ιδιοτήτων, αλλιώς την απορρίπτουμε.

Παράδειγμα

Υποθέτουμε ότι σε μια μελέτη οι μαθητές ταξινομούνται με βάση το χρώμα των μαλλιών και των ματιών τους (δύο χαρακτηριστικά - κριτήρια). Για 600 μαθητές καταγράψαμε τα ακόλουθα αποτελέσματα:

		Κατηγορία 2 (χρώμα μαλλιών)				
		Ξανθά 1	Καστανά 2	Μαύρα 3	Κόκκινα 4	Σύνολο
Κατηγορία 1 (χρώμα ματιών)	Μπλε 1	60 (33,3)	40 (100)	60 (33,3)	40 (33,3)	200 (200)
	Μαύρα 2	20 (16,7)	50 (50,0)	20 (16,7)	10 (16,7)	100 (100)
	Πράσινα 3	10 (16,7)	50 (50,0)	10 (16,7)	30 (16,7)	100 (100)
	Καφέ 4	10 (33,3)	160 (100,0)	10 (33,3)	20 (33,3)	200 (200)
	Σύνολο	100 (100)	300 (300)	100 (100)	100 (100)	600 (600)

Η υπόθεση H_0, η οποία θέλουμε να εξεταστεί σε επίπεδο $\alpha = 0{,}05$, είναι ότι τα δύο κριτήρια με βάση τα οποία ταξινομήθηκαν οι μαθητές είναι ανεξάρτητα μεταξύ τους. Η εναλλακτική υπόθεση H_1 είναι ότι τα δύο κριτήρια δεν είναι ανεξάρτητα. Μπορούμε να απορρίψουμε την H_0 με τα δεδομένα του πίνακα;

Απάντηση

Βήμα 1ο: Δηλώνουμε τις υποθέσεις και το επίπεδο α.

H_0: Τα κριτήρια είναι ανεξάρτητα. H_1: Τα κριτήρια δεν είναι ανεξάρτητα. Το επίπεδο σημαντικότητας είναι $\alpha = 0{,}05$.

Όπως και στην προηγούμενη παράγραφο, για να ελέγξουμε την H_0, πρέπει να συγκρίνουμε τις παρατηρούμενες (πραγματικές) συχνότητες με τις αναμενόμενες όταν η H_0 είναι αληθής. Αυτή η σύγκριση των f_o και f_e θα γίνει με το x^2 - κριτήριο.

Αν με R_i, $i = 1, 2, 3, 4$, συμβολίσουμε το άθροισμα των πραγματικών συχνοτήτων της i γραμμής (δηλαδή $R_1 = 200$, $R_2 = 100$, $R_3 = 100$ και $R_4 = 200$), με C_j, $j = 1, 2, 3, 4$, συμβολίσουμε το άθροισμα των πραγματικών συχνοτήτων στην j στήλη (δηλαδή $C_1 = 100$, $C_2 = 300$, $C_3 = 100$ και $C_4 = 100$), τότε οι αναμενόμενες συχνότητες στο κουτάκι (i, j) του πίνακα θα υπολογίζονται ως εξής:

$$E_{ij} = \frac{R_i C_j}{v}, \quad \text{όπου } v = 600$$

Οι υπολογισμοί αυτοί έγιναν και γράφονται μέσα στις παρενθέσεις σε κάθε κουτάκι. Για παράδειγμα, στο κουτάκι (1, 1) είναι $f_o = 60$ και $f_e = 33{,}3$. Η f_e υπολογίστηκε ως $E_{11} = \dfrac{200 \cdot 100}{600} = 33{,}3$. Στο κουτάκι (2, 3) είναι $f_o = 20$ και $f_e = 16{,}7$. Η f_e εδώ υπολογίστηκε ως $E_{23} = \dfrac{R_2 \cdot C_3}{600} = \dfrac{100 \cdot 100}{600} = 16{,}7$ κ.λπ.

Βήμα 2ο: Οι βαθμοί ελευθερίας είναι β.ε. $= (r-1)(c-1) = 3 \cdot 3 = 9$, οπότε $x^2_{0,05}(9) = 16{,}919$.

Βήμα 3ο: Προσδιορίζουμε τις αναμενόμενες συχνότητες όταν η H_0 είναι αληθής. Αυτές έχουν υπολογιστεί, όπως είπαμε προηγουμένως, και γράφτηκαν μέσα σε παρένθεση δίπλα στις πραγματικές συχνότητες. Αυτό το κάναμε για να αποφύγουμε την κατασκευή ενός νέου πίνακα με τις $f_{e_{ij}}$, αλλά και γιατί θα υπολογίζονται πιο γρήγορα οι διαφορές $f_{o_{ij}} - f_{e_{ij}}$.

Υπολογίζουμε τώρα το x^2 - στατιστικό από τον τύπο

$$x^2 = \sum \frac{(f_{o_{ij}} - f_{e_{ij}})^2}{f_{e_{ij}}}, \quad i, j = 1, 2, 3, 4$$

Αυτό είναι

$$x^2 = \frac{(60-33{,}3)^2}{33{,}3} + \frac{(40-100)^2}{100} + \frac{(60-33{,}3)^2}{33{,}3} + \frac{(40-33{,}3)^2}{33{,}3} +$$

$$+ \frac{(20-16{,}7)^2}{16{,}7} + \frac{(50-50)^2}{50} + \frac{(20-16{,}7)^2}{16{,}7} + \frac{(10-16{,}7)^2}{16{,}7} +$$

$$+ \frac{(10-16{,}7)^2}{16{,}7} + \frac{(50-50)^2}{50} + \frac{(10-16{,}7)^2}{16{,}7} + \frac{(30-16{,}7)^2}{16{,}7} +$$

$$+ \frac{(10-33{,}3)^2}{33{,}3} + \frac{(160-100)^2}{100} + \frac{(10-33{,}3)^2}{33{,}3} + \frac{(20-33{,}3)^2}{33{,}3} = 174$$

Βήμα 4ο: Η τιμή του στατιστικού $174 > 16{,}919$ πέφτει μέσα στην κρίσιμη περιοχή. Έτσι απορρίπτουμε την H_0 και ο ερευνητής μπορεί να συμπεράνει ότι το χρώμα των ματιών και το χρώμα των μαλλιών έχουν κάποια εξάρτηση μεταξύ τους.

Σημείωση: Τονίζουμε ότι αν σε κάποιο κουτάκι του πίνακα παρουσιαστεί αναμενόμενη συχνότητα μικρότερη του 5, τότε αυτή πρέπει να προστεθεί σε κάποιο άλλο κουτάκι ώστε να ικανοποιείται ο περιορισμός που επιβάλλει να είναι οι αναμενόμενες συχνότητες μεγαλύτερες του 5. Τότε όμως θα ελαττωθούν και οι βαθμοί ελευθερίας, αφού θα ελαττωθεί το πλήθος των γραμμών ή στηλών.

6.4 Εισαγωγή στην ανάλυση της διασποράς (ANOVA)

- *Γενικά*

Στην πράξη πολλές φορές παρουσιάζονται προβλήματα όπου πρέπει να συγκριθούν οι μέσες τιμές για περισσότερα από δύο δείγματα.

Για την αντιμετώπιση των προβλημάτων αυτών αναπτύχθηκε στη Στατιστική μια μέθοδος που ονομάζεται «**ανάλυση της διασποράς** ή **ανάλυση της διακύμανσης**».

Με απλά λόγια, η ανάλυση της διασποράς είναι μια διαδικασία ελέγχου υποθέσεων που τη χρησιμοποιούμε για να εκτιμήσουμε τις διαφορές μεταξύ δύο ή περισσότερων πληθυσμών (θεραπειών). Όπως όλες οι άλλες συμπερασματικές διαδικασίες έτσι και αυτή χρησιμοποιεί ως βάση δεδομένα δείγματος προκειμένου να εξαγάγει συμπεράσματα για τους πληθυσμούς. Συχνά χρησιμοποιούμε t - τεστ για να ελέγξουμε τη διαφορά των μέσων τιμών. Το t - τεστ μας βοηθά μόνο για δύο πληθυσμούς. Για περισσότερους πληθυσμούς χρησιμοποιούμε τη νέα μέθοδο που αναφέραμε, την ανάλυση της διασποράς. Γι' αυτό η ανάλυση της διασποράς παρέχει στους ερευνητές μεγαλύτερη ευελιξία στον σχεδιασμό πειραμάτων και στην ερμηνεία αποτελεσμάτων.

Όπως και το t - τεστ έτσι και η ανάλυση της διασποράς μπορεί να χρησιμοποιηθεί είτε σε διαφορετικά δείγματα (ένα από κάθε πληθυσμό) είτε στο ίδιο δείγμα, εφαρμόζοντας διαφορετικές προσεγγίσεις - πειράματα (θεραπείες, διδακτικές μεθόδους κ.λπ.). Επιπρόσθετα, η ανάλυση της διασποράς μπορεί να χρησιμοποιηθεί στην εκτίμηση των αποτελεσμάτων μιας έρευνας στην οποία εμπλέκονται περισσότερες από μία ανεξάρτητες μεταβλητές. Για παράδειγμα, ένας εκπαιδευτικός είναι δυνατό να θελήσει να συγκρίνει την αποτελεσματικότητα δύο διαφορετικών διδακτικών μεθόδων (πρώτη ανεξάρτητη μεταβλητή) σε τρεις τάξεις διαφορετικών μεγεθών (δεύτερη ανεξάρτητη μεταβλητή). Η εξαρτημένη μεταβλητή εδώ είναι ο βαθμός κάθε μαθητή σ' ένα σταθμισμένο τεστ. Η δομή αυτού του πειράματος, το οποίο περιλαμβάνει τη σύγκριση δειγματικών μέσων για έξι διαφορετικές περιπτώσεις, έχει ως εξής:

Ανεξάρτητη μεταβλητή 2: Μέγεθος τάξης

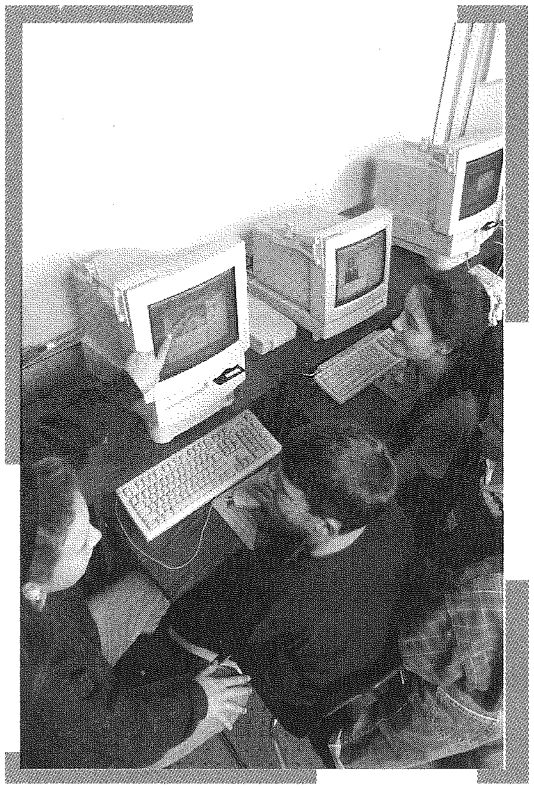
		Μικρή τάξη	Μεσαία τάξη	Μεγάλη τάξη
Ανεξάρτητη μεταβλητή 1: Διδακτικές μέθοδοι	Μέθοδος Α	Δείγμα 1	Δείγμα 2	Δείγμα 3
	Μέθοδος Β	Δείγμα 4	Δείγμα 5	Δείγμα 6

Όπως βλέπουμε, η ανάλυση της διασποράς παρέχει στους ερευνητές μια εξαιρετικά ευέλικτη τεχνική ανάλυσης δεδομένων. Μπορεί να χρησιμοποιηθεί στην εκτίμηση της σημαντικότητας των διαφορών των μέσων τιμών σε μια ποικιλία ερευνητικών καταστάσεων και είναι η διαδικασία που χρησιμοποιείται περισσότερο στον έλεγχο υποθέσεων.

Δημήτρης Λ. Καραγεώργος
ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΑΘΗΝΩΝ
Επιτ. ΣΥΜΒΟΥΛΟΣ ΤΟΥ ΠΑΙΔΑΓΩΓΙΚΟΥ ΙΝΣΤΙΤΟΥΤΟΥ

ΜΕΘΟΔΟΛΟΓΙΑ ΕΡΕΥΝΑΣ στις επιστήμες της αγωγής

ΜΙΑ ΔΙΔΑΚΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ



Σαββάλας
ΕΚΔΟΣΕΙΣ